The sign problem and the Lefschetz thimble

Marco Cristoforetti, Luigi Scorzato[1]
ECT*. Strada delle tabarelle, 286 – I-38123, Trento, Italy.

Francesco Di Renzo
Università di Parma and INFN gruppo collegato di Parma. Viale G.P. Usberti n.7/A – I-43124, Parma, Italy.

Abstract. Recently, we have introduced a novel approach to deal with the sign problem that prevents the Monte Carlo simulations of a class of quantum field theories (QFTs). The idea is to formulate the QFT on a Lefschetz thimble. Here we review the formulation of our approach and describe the *Aurora* Monte Carlo algorithm that we are currently testing on a scalar field theory with a sign problem.

1. Introduction
Recently, we have proposed a novel approach [1] to deal with the sign problem that hinders Monte Carlo simulations of many quantum field theories (QFTs). The approach consists in formulating the QFT on a Lefschetz thimble. In this paper we concentrate on the application to a scalar field theory with a sign problem. In particular, we review the formulation and the justification of the approach, and we also describe the *Aurora* Monte Carlo algorithm that we are currently testing.

2. Formulation
In this paper we consider a scalar field theory with $U(1)$ global symmetry:

$$S = \int d^4x[|\partial\phi|^2 + (m^2 - \mu^2)|\phi|^2 + \mu j_0 + \lambda|\phi|^4], \qquad j_\nu := \phi^* \overleftrightarrow{\partial_\nu} \phi, \tag{1}$$

where $\phi(x)$ is a complex scalar field. This model can be regularized on a lattice as:

$$S[\phi] = \sum_x \left[\left(2d + m^2\right) \phi_x^* \phi_x + \lambda(\phi_x^* \phi_x)^2 - \sum_{\nu=0}^{d-1} \left(\phi_x^* e^{-\mu\delta_{\nu,0}} \phi_{x+\hat\nu} + \phi_{x+\hat\nu}^* e^{\mu\delta_{\nu,0}} \phi_x \right) \right],$$

and has a sign problem when $\mu \neq 0$, which has been successfully treated in [2] and also in [3].

Our approach consists in defining the observables as:

$$\langle\mathcal{O}\rangle_0 = \frac{1}{Z_0} \int_{\mathcal{J}_0} \prod_{a,x} d\phi_{a,x}\, e^{-S[\phi]} \mathcal{O}[\phi], \qquad Z_0 = \int_{\mathcal{J}_0} \prod_{a,x} d\phi_{a,x}\, e^{-S[\phi]}, \tag{2}$$

[1] Speaker

where the integration domain $\mathcal{J}_0$ is the Lefschetz thimble [4, 5] attached to a global minimum ϕ_{glob} of the real part of the action $S_R = \Re S$, when restricted to the original domain $\mathcal{C} = \mathbb{R}^{2V}$. More precisely, $\mathcal{J}_0$ is the manifold of real dimension $n = 2V$, defined as union of all the curves of steepest descent (SD) for S_R, i.e., solutions of:

$$\begin{aligned} \frac{d}{d\tau}\phi^{(R)}_{a,x}(\tau) &= -\frac{\delta S_R[\phi(\tau)]}{\delta\phi^{(R)}_{a,x}}, \quad \forall a, x, \\ \frac{d}{d\tau}\phi^{(I)}_{a,x}(\tau) &= -\frac{\delta S_R[\phi(\tau)]}{\delta\phi^{(I)}_{a,x}}, \quad \forall a, x, \end{aligned} \tag{3}$$

that end in ϕ_{glob} for $\tau \to \infty$. Here we have assumed the usual complexification of the complex scalar field (see, e.g., [3]), where both the real (ϕ_1) and imaginary (ϕ_2) part of the original fields become complex $\phi_a = \phi_a^{(R)} + i\phi_a^{(I)}$, $a = 1, 2$.

2.1. Formulation in presence of SSB

Note that the formulation described above is possible only when ϕ_{glob} is a minimum of the action with positive definite Hessian. In presence of symmetries that act non-trivially in ϕ_{glob} this is not the case.

Gauge symmetries can be treated as described in [6, 5, 1], where, essentially, the thimble is defined modulo gauge transformations. But this is not suitable to study the possibility of spontaneous symmetry breaking (SSB). However, the proper way to study SSB (even in the ordinary formulation of the functional integral) is to introduce a small term of explicit symmetry breaking and study the limit when such term goes to zero. In this way, ϕ_{glob} becomes a true minimum and the thimble can be defined.

2.2. Justification of the approach

The functional integral in Eq. (2) does not coincide, in general, with the standard formulation. Morse theory [5] only enables us to say that—under suitable conditions on $S[\phi]$ and $\mathcal{O}$— the standard functional integral coincides with an integral over a particular combination of thimbles $\sum_\sigma n_\sigma \mathcal{J}_\sigma$, for some integer n_σ, and where the sum runs over all the stationary points of the complexified action[2].

However, we have shown in [1] that the model defined by Eq. (2) has the same degrees of freedoms, the same symmetries and symmetry representations, and also the same perturbative expansion as the standard formulation. In this sense, by universality, the formulation of Eq. (2) may be seen as an alternative regularization of the model in Eq. (1).

3. The *Aurora* algorithm

In this section we describe an updated form of the algorithm proposed in [1]. What we want to compute are expectation values like:

$$\frac{1}{Z_0} e^{-iS_I} \int_{\mathcal{J}_0} \prod_x d\phi_x \; e^{-S_R[\phi]} \mathcal{O}[\phi], \tag{4}$$

where the imaginary part of the action S_I is constant in $\mathcal{J}_0$ and can be taken out of the integral. Hence the phase of e^{-S} cancels from all observables. Moreover, the real part S_R is bounded from below by its value at the stationary point[3]. The functional integral in Eq. (4) possesses

[2] An argument in [5] (see Sec. 3.2), can be used to show that only the stationary points near to the global minima in $\mathcal{C}$ can bring a non negligible contribution. See [1], Sec. II.B.3.

[3] The integral is convergent, because $S[\phi]$ is a polynomial in ϕ and $\mathcal{J}_0$ is union of the curves of SD.

a bounded real action and hence can be studied, e.g., with a Langevin algorithm[4]. The choice of the Langevin algorithm is particularly convenient because the drift term of the Langevin equation for the real part of the action coincides with Eq. (3), i.e., the SD. Hence, such drift preserves the manifold $\mathcal{J}_0$ by construction. The Langevin noise, however, needs to be projected on the tangent space to $\mathcal{J}_0$, in order to explore the integration domain correctly.

The projection on the tangent space of $\mathcal{J}_0$ at ϕ is challenging, because we seem to have no way to tell which configurations in the neighborhood of $\phi \in \mathcal{J}_0$ will eventually also fall in ϕ_{glob}, by following the SD flow. However, the tangent space at the stationary point ϕ_{glob} can be computed, by computing the Hessian matrix[5]. In order to use this fact to generate a noise vector tangent to $\mathcal{J}_0$ at ϕ, we need a way to transport a vector η along the flow of SD, while keeping it tangent to $\mathcal{J}_0$. One way to achieve this is by requesting that the vector η is parallel transported along the flow of S_R, i.e. requesting that its Lie derivative along ∂S_R is zero:

$$\mathcal{L}_{\partial S_R}(\eta) = [\partial S_R, \eta] = 0$$

which nicely translates into a first order ordinary differential equation, which is also linear in η:

$$\frac{d}{d\tau}\eta_j(\tau) = \sum_k \eta_k(\tau)\partial_k\partial_j S_R. \tag{5}$$

Note that we use j, k as multi-indices for $(R/I, a, x)$.

Actually, we can do better. Eq. (5) ensures a parallel transport of η along the flow ∂S_R. This is actually more than what we need, because we only need to keep η tangent to $\mathcal{J}_0$. On the other hand, it would be nice to transport η by means of an orthogonal flow, that preserves orthonormality, isotropy, and also enables the use of stable numerical integrators (see Sec. 5.3 of [7]). This can be done quite easily by employing the Iwasawa decomposition of $\partial^2 S_R$ (see, e.g., Sec. VI of [8]), which, at the group level, is a generalization of the Gram-Schmidt orthonormalization procedure, and it can be implemented also at the algebra level. More precisely, if we write:

$$\left(\partial^2 S_R\right)_{i,j} = A_{i,j} + D_{i,j} + N_{i,j},$$

where A is skew-symmetric, D is diagonal and N is upper triangular, then the flow

$$\frac{d}{d\tau}\eta(\tau) = \eta(\tau)A(\tau) \tag{6}$$

defines an orthogonal flow whose solutions η span the same space spanned by the solutions of Eq. (5). This means that the tangent space of $\mathcal{J}_0$ is still preserved. But now, the flow defined by Eq. (6) also preserves orthonormality, isotropy, and such properties can be even ensured numerically by using suitable numerical integrators, such as the implicit midpoint rule [7].

We can now summarize the algorithm as follows. We use t to represent the Langevin (Monte Carlo) time and τ to represent the parameter of the SD. Let us assume that we have already a configuration ϕ_t in $\mathcal{J}_0$, and we also have a path $\phi_t(\tau)$ that fulfills the equations of SD and connects ϕ_t to a configuration $\phi_t^{(\varepsilon)}$ with norm less than ε. The value of ε must be sufficiently small so that the quadratic approximation of the action is valid. One step of Langevin proceeds as follows:

- apply the Langevin force, that consists in taking one step forward along the curve of SD which is already available. The result is the configuration ϕ'_t;

[4] The phase that comes from the measure $d\phi$ will be treated with reweighting, as discussed in Sec. 3.1.

[5] Which is easy to compute also analytically, when $\phi_{\text{glob}} = 0$, and can be computed numerically, once and for all, for general ϕ_{glob}.

- extract a Gaussian noise η;
- project out the components of η orthogonal to $\mathcal{J}_0$ at the origin. More precisely, set $\eta_{\|} = (P-1)\eta$, where $P = \partial^2 S_R/\sqrt{(\partial^2 S_R)}$, and then re-scale the norm of $\eta_{\|}$ according to the χ distribution;
- evolve $\eta_{\|}$ with Eq. (6) from $\phi_t^{(\varepsilon)}$ to ϕ'_t. The orthogonal properties of the flow can also be preserved numerically by employing, e.g., the implicit midpoint rule. This leads to solve a 5-D linear system in $\eta_{\|}(\tau)$;
- define a tentative new configuration as $\phi_{t+dt}^{\text{guess}} = \phi'_t + \sqrt{2t}\,\eta(\tau)$;
- use $\phi_{t+dt}^{\text{guess}}$ as a starting point to find a solution of the SD equation. This can also be integrated with the implicit midpoint rule with the constraint that $(P-1)\phi_{t+dt}^{(\varepsilon)} = 0$. This leads to a non-linear 5-D system that can be solved—if the initial guess is sufficiently good—with a number of Newton-Raphson iterations.

3.1. The residual phase

As noted in the previous section, the integral of Eq. (4) also includes a residual phase that comes from the determinant of the tangent space of $\mathcal{J}_0$ at ϕ, and is not necessarily real and positive. We argue that this should not represent a *sign problem*. In fact, a sign problem means that the average phase $|\langle\prod d\phi\rangle| \ll 1$. However, that phase is constant in the region where the integrand is not very small. In fact, such phase is completely neglected in the saddle point expansion, which, on the other hand, can be identified with some form of perturbative expansion of the QFT. Saying that the average phase could be negligible is equivalent to saying that the "perturbative contribution" (in some sense that we cannot define precisely) could be negligible. We do expect that a non-perturbative contribution may be important, but that the perturbative one could be negligible is contrary to the general expectations.

In this sense, we have good reasons to hope that residual phase can be treated with reweighting, but we have no justification to neglect it. If so, we have to compute it, which is quite expensive. This can be done as follow. If we call T_ϕ the complex $(2V \times 2V)$ matrix that spans the tangent space of $\mathcal{J}_0$ at ϕ, we have:

$$\det(T_{\phi_\tau}) = -\int_\tau^\infty ds\,\mathrm{Tr}[\,T_{\phi_s}^{-1}\,\frac{d}{ds}T_{\phi_s}\,] \simeq \sum_{i=0}^{N_\tau}\mathrm{Tr}[\,T_{\phi_i}^{-1}\,\Delta T_{\phi_i}\,] = \sum_{i=0}^{N_\tau}\sum_{k=1}^{N_R}{\xi^{(k)}}^T T_{\phi_i}^{-1}\,\Delta T_{\phi_i}\,\xi^{(k)}$$

where the $\xi^{(k)}$, $k = 1,\ldots,N_R$ are noisy estimators for the trace. Assuming that the inversion of T_ϕ would require N^{CG} iterations, the computation of the determinant would cost as much as evolving $N_R N^{CG}$ vectors, which is described in the previous section.

Some saving may also come from the fact that the computation of the phase is necessary only for the configurations that are actually used for measurement.

3.2. Preliminary tests

As a first test, we implemented the method by using the naive Euler integration method. Although bound to fail because of the nature of the Euler method, this is an interesting test in order to get a feeling of how difficult it is to keep the system on a thimble and how sensitive the system is to perturbations. Some results are shown in Fig. 1 and 2. These are encouraging, because it seems that even the Euler method is able to keep the system on the thimble long enough to see a thermalization. This lets us hope that a better integrator will have good chances to converge. Observables appear around the expected order of magnitude, but their thermalization is not clear.

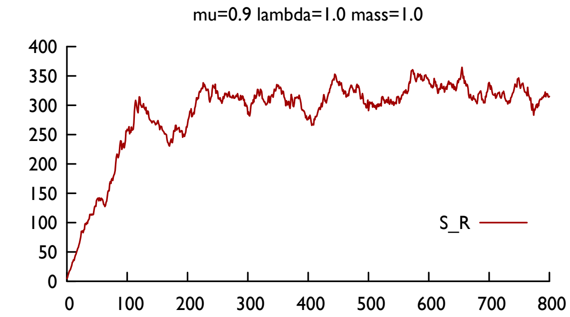

Figure 1. The real part of the action on a $V = 4^4$ lattice with Euler integrator. The system is stable long enough to show signs of a thermalization.

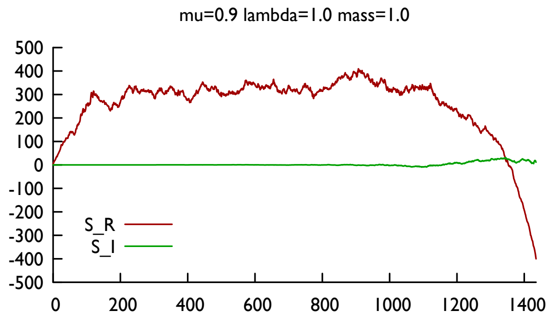

Figure 2. Same as Fig. 1, but for a longer time. The Euler integrator cannot constrain the system on the thimble, and the system drifts away at some point. This is also confirmed by the deviation of the imaginary part of the action from a constant.

4. Conclusions

We have illustrated an new approach to deal with the sign problem that afflicts a class of QFTs. It consists in regularizing the QFT on a Lefschetz thimble. Although it does not coincide with the usual regularization, it is a legitimate one on the basis of universality. In fact, we could prove that QCD on the thimble has the same symmetries, d.o.f. and PT as the usual formulation.

We have also introduced a Monte Carlo algorithm to achieve an importance sampling of the configurations on the thimble. Its numerical implementation will be certainly challenging and expensive, but all the steps of the algorithm are, a priori, feasible and have acceptable scaling. The residual phase should not give a sign problem (unless we believe that the perturbative contribution can be negligible) and hence should be manageable with reweighting, but this must be checked. We are presently testing the method for a scalar QFT on tiny lattices with encouraging results.

Acknowledgments

This research is supported by the AuroraScience project (funded by the Provincia Autonoma di Trento and INFN), and by the Research Executive Agency (REA) of the European Union under Grant Agreement No. PITN-GA-2009-238353 (ITN STRONGnet). L.S. and M.C. are members of LISC. FDR is partially supported by INFN i.s. MI11 and by MIUR contract PRIN2009 (20093BMNPR_004).

References

[1] Cristoforetti M, Di Renzo F and Scorzato L (AuroraScience Collaboration) 2012 *Phys. Rev. D* **86**(7) 074506 (*Preprint* `1205.3996`)
[2] Endres M G 2007 *Phys.Rev.* **D75** 065012 (*Preprint* `hep-lat/0610029`)
[3] Aarts G 2009 *Phys.Rev.Lett.* **102** 131601 (*Preprint* `0810.2089`)
[4] Pham F 1983 *Proc. Symp. Pure Math.* **40**(2) 319–333
[5] Witten E 2010 (*Preprint* `1001.2933`)
[6] Atiyah M and Bott R 1983 *Phil.Trans.Roy.Soc.Lond.* **A308** 523–615
[7] Iserles A 2008 *A First Course in the Numerical Analysis of Differential Equations* 2nd ed Cambridge Texts in Applied Mathematics (Cambridge University Press) ISBN 9780521734905
[8] Knapp A 2002 *Lie Groups Beyond an Introduction* 2nd ed Progress in Mathematics (Birkhäuser Boston) ISBN 9780817642594